

ABSTRACT OF THE DISCLOSURE

A real-time networked telephony or computer system has a feature complex and/or applications that offer a class of features to a subscriber, including call information, and permits the subscriber to manage incoming and existing calls through available features accessed using spoken utterances. A speech processing unit coupled to the system interprets a subscriber's spoken utterances without requiring the subscriber to train the system to recognize his or her voice. The interpretation of spoken utterances is enabled by a system state database that is maintained at the speech processing unit and comprises a database of the possible system states, including possible call flows for a call, and a database associated with the system state database comprising context-specific grammar that a subscriber may recite at respective points in the call flow. The speech processing unit may also convert message signals from the network to speech which is read to the subscriber using a text to speech translator. The network can identify the voice or subscriber voice, or language used and will thereafter recognize all further commands using specific grammar for that language as well as perform text-to-speech conversion using the identified language. Use of the features can be applied to update of grammars, profiles and templates, *etc.* by transmitting results of transactions.